

GALLAGER-symmetrisch schwach-symmetrisch stark symmetrisch (Zeilen sind Permutationen voneinander) (somit GALLAGER- und schwach-symmetrisch)

25

Übung 0.1. Die folgenden Kanalmatrizen

$$\begin{pmatrix} 1/8 & 3/8 & 3/8 & 1/8 \\ 3/8 & 3/8 & 3/8 & 1/8 \end{pmatrix}$$

ist stark symmetrisch.

$$\begin{pmatrix} 2/3 & 1/3 & 0 \\ 0 & 1/3 & 2/3 \end{pmatrix}$$

schwach und GALLAGER-symmetrisch, aber nicht stark.

$$\begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.2 & 0.1 & 0.7 \end{pmatrix}$$

hat keine Symmetrie.

$$\begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.2 & 0.1 & 0.4 & 0.3 \end{pmatrix}$$

GALLAGER-symmetrisch.

$$\begin{pmatrix} 0.3 & 0.2 & 0.5 \\ 0.2 & 0.5 & 0.3 \\ 0.5 & 0.3 & 0.2 \end{pmatrix}$$

stark symmetrisch.

$$\begin{pmatrix} p & 1-p & 0 & 0 \\ 1-p & p & 0 & 0 \\ 0 & 0 & q & 1-q \\ 0 & 0 & 1-q & q \end{pmatrix}$$

ist stark symmetrisch für $p = 1 - q$ sonst nichts.

$$\begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.2 & 0.4 & 0.3 & 0.1 \\ 0.3 & 0.1 & 0.2 & 0.4 \\ 0.4 & 0.3 & 0.2 & 0.1 \end{pmatrix}$$

ist nur schwach symmetrisch.

26

Übung 0.2.

Lösung. Gegeben ist eine Kanalfolge K_1, K_2 mit X_1 am Eingang und Y_1 am Ausgang.

(a)

$$C_1 = \max_{p_{X_1}} I(X_1, Y_1)|_{p_{X_1}},$$

dabei sei $p_{X_1}^*$ die optimale Verteilung zu Kanal 1.

$$C = I(X_1, Y_2)|_{p_{X_1}^*},$$

mit $p_{X_1}^*$ die optimale Verteilung für den Gesamtkanal. Nun gilt laut Datenverarbeitungsungleichung

$$C = I(X_1, Y_2)|_{p_{X_1}^*} \leq I(X_1, Y_1)|_{p_{X_1}^*} \leq I(X_1, Y_1)|_{p_{X_1}^{\prime*}} = C_1$$

(b) Die Kanalmatrix vom Gesamtkanal ermittelt sich zu

$$K = K_1 K_2.$$

(c) Gegeben sind zwei symmetrische Binärkanäle (BSC) mit Fehlerwahrscheinlichkeit ε_1 , Damit

$$K_1 = \begin{pmatrix} 1 - \varepsilon_1 & \varepsilon_1 \\ \varepsilon_1 & 1 - \varepsilon_1 \end{pmatrix}$$

analog ergibt sich

$$K_2 = \begin{pmatrix} 1 - \varepsilon_2 & \varepsilon_2 \\ \varepsilon_2 & 1 - \varepsilon_2 \end{pmatrix}.$$

Die Kanalmatrix des Gesamtkanals eine Matrix

$$K = K_1 K_2 = \begin{pmatrix} 1 - \varepsilon & \varepsilon \\ \varepsilon & 1 - \varepsilon \end{pmatrix}$$

mit $\varepsilon = \varepsilon_1(1 - \varepsilon_2) + \varepsilon_2(1 - \varepsilon_1)$. D.h. die Gesamtkanalkapazität ist BSC mit Fehlerwahrscheinlichkeit ε . Also laut Script (3.19)

$$C = 1 - H_b(\varepsilon).$$

(d) Wieder wird ein symmetrischer Binärkanal K_1 (BSC) mit einem Binärkanal mit Auslöschung (BEC) hintereinandergeschaltet:

$$K_1 = \begin{pmatrix} 1 - \varepsilon_2 & \varepsilon_2 & 0 \\ 0 & \varepsilon_2 & 1 - \varepsilon_2 \end{pmatrix}$$

Matrixmultiplikation liefert dann

$$K = \begin{pmatrix} (1 - \varepsilon_2)(1 - \varepsilon_1) & \varepsilon_2 & \varepsilon_1(1 - \varepsilon_2) \\ \varepsilon_1(1 - \varepsilon_2) & \varepsilon_2 & (1 - \varepsilon_2)(1 - \varepsilon_1) \end{pmatrix}.$$

Diese ist GALLAGER-symmetrisch, denn die Zeilen sind Permutationen voneinander und es existiert eine Zerlegung in stark symmetrische Matrizen. Die optimale Eingangsverteilung ist also die Gleichverteilung. Es

berechnet sich dann $I(X_1, Y_2) = H(Y_2) - H(Y_2|X_1)$ für X_1 gleichverteilt. Es ergibt sich

$$\begin{array}{cccc} Y_1 & 0 & \Delta & 1 \\ \hline p_{Y_2}(y_2) & \frac{1}{2}(1 - \varepsilon_2) & \varepsilon_2 & \frac{1}{2}(1 - \varepsilon_2) \end{array}$$

also $H(Y_2) = (1 - \varepsilon_2) + H_b(\varepsilon_2)$ und $H(Y_2|Y_1) = \frac{1}{2}H(Y_2|X_1 = 0) + \frac{1}{2}H(Y_2|X_1 = 1) = (1 - \varepsilon_2)H_b(\varepsilon_1) + H_b(\varepsilon_2)$. Also

$$C = (1 - \varepsilon_2)(1 - H_b(\varepsilon_1)).$$

Bei $\varepsilon_1 = 0$ bleibt nur der BEC über, bei $\varepsilon_2 = 0$ nur der BSC.

27

Lösung.

- (a) Die Kanalmatrix für den Gesamtkanal ist zu ermitteln. Diese ergibt sich zu

$$K = \alpha K_1 + (1 - \alpha) K_2,$$

da mit Wahrscheinlichkeit α der Kanal K_1 und mit $1 - \alpha$ Kanal K_2 durchlaufen wird.

- (b) Für feste Eingangsverteilung ist die Transinformation I konvex bzgl. der Kanalmatrix (1.24), d.h.

$$I(p_X, \alpha K_1 + (1 - \alpha) K_2) \leq \alpha I(p_X, K_1) + (1 - \alpha) I(p_X, K_2).$$

mit $I(X, Y) =: I(p_X, K)$. Sei $p_X^{(1)*}$ optimale Verteilung zu K_1 und $p_X^{(2)*}$ optimale Verteilung zu K_2 , sowie p^* optimale Verteilung zu K . Daraus folgt

$$\begin{aligned} C &= I(p_X^*) \leq \alpha I(p_X^*, K_1) + (1 - \alpha) I(p_X^*, K_2) \\ &\leq \alpha I(p_X^{(1)*}, K_1) + (1 - \alpha) I(p_X^{(2)*}, K_2) \\ &= \alpha C_1 + (1 - \alpha) C_2, \end{aligned}$$

was die gesuchte Identität ist nach Definition der Informationskapazität.

- (c) Kanalmatrix für den Gesamtkanal (wieder zwei BSC's hintereinandergeschaltet) ergibt sich zu

$$K = \alpha \begin{pmatrix} 1 - \varepsilon_1 & \varepsilon_1 \\ \varepsilon_1 & 1 - \varepsilon_1 \end{pmatrix} + (1 - \alpha) \begin{pmatrix} 1 - \varepsilon_2 & \varepsilon_2 \\ \varepsilon_2 & 1 - \varepsilon_2 \end{pmatrix} = \begin{pmatrix} 1 - \varepsilon & \varepsilon \\ \varepsilon & 1 - \varepsilon \end{pmatrix}$$

mit $\varepsilon = \alpha \varepsilon_1 + (1 - \alpha) \varepsilon_2$. D.h. der Gesamtkanal ist BSC mit Fehlerwahrscheinlichkeit ε , damit laut Script $C = 1 - H_b(\varepsilon)$.

1 4. Hausaufgabe zur LV Informationstheorie (28-30)

19.06.2014

Übung 1.1 (Stark- und schwach typische Sequenzen diskreter gedächtnisloser Quellen). Gegeben ist eine diskrete Gedächtnislose Quelle $(U_k)_k$ mit einem binären Alphabet $\mathcal{U} = \{0, 1\}$. Die Auftretswahrscheinlichkeit für das Symbol 0 sei q .

$$s_1 = (00100)$$

$$s_2 = (10101)$$

$$s_3 = (10111)$$

$$s_4 = (11111)$$

Entscheiden Sie für jede dieser Sequenzen, ob die Sequenz ε -stark und oder ε -schwach-typisch ist, wenn ε und q folgende Werte haben.

(a) $\varepsilon = 0,15$, $q = 1/3$.

(b) $\varepsilon = 0,3$, $q = 1/7$.

Bemerkung 1. Die Sequenz ist *schwach typisch*, wenn sie die Ungleichung $\left| -\frac{1}{n} \log_2(p_U^{(n)}(u^{(n)})) - H(U) \right| \leq \varepsilon$ erfüllt, der erste Term heißt dabei empirische Entropie (Bezeichnung $H(\hat{U})^{(n)}$). Ist die Sequenz stark typisch, so muss die Ungleichung $\left| \frac{1}{n} N(a|u^{(n)}) - p_U(a) \right| < \varepsilon/|U|$ für alle $a \in \mathcal{U}$.

Lösung. Es gilt $p_U(s_1) = q^3(1-q)^2$, $p_U(s_2) = q^2(1-q)^3$, $p_U(s_3) = q(1-q)^4$, $p_U(s_4) = (1-q)^5$.

(a) $\varepsilon = 0,15$, $q = 1/3$, $p_U(s_1) = 4/243$, $p_U(s_2) = 8/243$, $p_U(s_3) = 16/243$, $p_U(s_4) = 32/243$. Damit erhält man $\hat{H}(s_1) = 1,185\text{bit}$, $\hat{H}(s_2) = 0,985\text{bit}$, $\hat{H}(s_3) = 0,785\text{bit}$ und $\hat{H}(s_4) =$. Es ergibt sich also:

$$\left| \hat{H}(s_1) - H(U) \right| = 0,267$$

$$\left| \hat{H}(s_2) - H(U) \right| = 0,067$$

$$\left| \hat{H}(s_3) - H(U) \right| = 0,133$$

$$\left| \hat{H}(s_4) - H(U) \right| = 0,333$$

Damit ist der Abstand für s_2, s_3 kleiner als $\varepsilon = 0,15$ also sind sie schwach

typisch. Nun testen wir auf start-typisch.

s_i	$N(0 s_i)$	$N(1 s_i)$	$ 1/nN(a s_i) - p_U(a) $
s_1	3	2	$ 3/5 - 1/3 = 0,2667$
s_2	2	3	$ 2/5 - 1/3 = 0,067$
s_3	1	4	$ 1/5 - 1/3 = 0,133$
s_4	0	5	$1/3$

Damit sind s_2, s_3 stark- und und schwach-typisch. s_1, s_4 sind weder das eine noch das andere.

Bemerkung 2. Im binären gilt $|1/nN(0|s) - p_U(0)| = |1/nN(1|s) - p_U(1)|$, da $|1/nN(1|s) - p_U(1)| = |1/n(n - N(1|s)) - (1 - p_U(0))| = |1/nN(0|s) - p_U(0)|$.

- (b) Analog berechnet man, dass hier s_3 stark und schwach typisch ist, s_4 nur stark typisch ist und s_1, s_2 beides nicht sind.

Bemerkung 3. Es sind also alle Kombinationen aus stark und schwach typisch möglich.

Übung 1.2 (Informationskapazität). Gegeben seien zwei unabhängige Zufallsvariablen X und Z mit Alphabeten $\mathcal{X} = \{1, 2, 3\}$. Die Zufallsgröße Z besitzt die Wahrscheinlichkeitsdichte p_Z , die wie folgt gegeben ist:

z	1	2	3
$p_z(\delta)$	ε	δ	$1 - \delta - \varepsilon$

Es soll $\varepsilon, \delta \in [0, 1/2]$ gelten. Es sei Y durch $Y := X + Z \bmod 3$ gegeben.

- Gib das Alphabet von Y an.
- Bestimme die Wahrscheinlichkeiten $p_{Y|X}(\bullet|x)$ für alle $x \in \mathcal{X}$.
- Wir wollen nun X und Y als Eingang bzw. Ausgang eines DMC betrachten. Berechnen Sie die Informationskapazität dieser DMC in Abhängigkeit von ε, δ und geben Sie die zugehörige optimale Wahrscheinlichkeitsverteilung an.
- Sei $\varepsilon = \delta$, für welches ε nimmt nun die in der vorherigen Teilaufgabe berechnete Informationskapazität ihr Minimum und Maximum an.

Lösung.

(a) $\mathcal{Y} = \mathcal{X}$.

(b) Folgende Tabelle ergibt sich:

X	Z	$X + Z$	Y	$P_{Y X}(y x)$
1	1	2	2	δ
1	2	3	2	ε
1	3	4	1	$1 - \delta - \varepsilon$
2	1	3	0	δ
2	2	4	1	ε
2	3	5	2	$1 - \delta - \varepsilon$
3	1	4	1	δ
3	2	5	2	ε
3	3	6	0	$1 - \delta - \varepsilon$

Man hat also die Kanalmatrix:

$$p_{Y|X}(x, y) \begin{array}{ccc} & 0 & 1 & 2 \\ \begin{array}{c} 1 \\ 2 \\ 3 \end{array} & \begin{array}{c} \varepsilon \\ \delta \\ 1 - \delta - \varepsilon \end{array} & \begin{array}{c} 1 - \delta - \varepsilon \\ \varepsilon \\ \delta \end{array} & \begin{array}{c} \delta \\ 1 - \delta - \varepsilon \\ \varepsilon \end{array} \end{array}$$

(c) Der DMC ist symmetrisch. Die Informationskapazität wird durch Gleichverteilung am Eingang erreicht (Skript §3.21). Für die Kapazität gilt (§3.22)

$$C = \log_2 |\mathcal{Y}| - H(r)$$

hier also:

$$C = \log_2 3 - (-\varepsilon \log_2 \varepsilon - \delta \log \delta - (1 - \delta - \varepsilon) \log_2 (1 - \delta - \varepsilon)).$$

(d) Hier $\varepsilon = \delta$. Dann wird die Entropie maximal bei Gleichverteilung, die Kapazität also minimal für $\varepsilon = \delta = 1/3$. $C_{\min} = 0$. Die Entropie wird minimal für $\varepsilon = 0$, dann ist $C_{\max} = \log_2 3$.

Übung 1.3 (Hintereinanderschaltung von Multiplikationskanälen). Gegeben sei ein diskreter gedächtnisloser Multiplikationskanal mit binärem Ein- und Ausgang (X_1, Y_1) . Das multiplikative Rauschen wird durch die vom Kanaleingang unabhängige zufallsgröße Z_1 mit binärem Alphabet $\mathcal{Z}_1 = \{0, 1\}$ beschrieben, wobei $P(Z_1 = 1) = \varepsilon_1$ sei.

(a) Bestimme die Kanalmatrix und die Informationskapazität.

Lösung.

(a) Siehe Aufgabe 16:

$$K_1 = \begin{pmatrix} 1 & 0 \\ 1 - \varepsilon_1 & \varepsilon_1 \end{pmatrix}$$

Es gilt $C_1 = \max_{p_X} I(x; y_1) = \log_2(1 + \varepsilon_1(1 - \varepsilon_1)^{\frac{1-\varepsilon_1}{\varepsilon_1}})$ bit mit

$$p_X = \frac{1}{(1 - \varepsilon_1)^{\frac{\varepsilon_1-1}{\varepsilon_1}} + \varepsilon_1}.$$

(b) Durch Hintereinanderschaltung zweier Kanäle entsteht die Gesamtkanalmatrix durch Multiplikation der beiden Teilkanalmatrizen ($K = K_1 K_2$).

$$K = \begin{pmatrix} 1 & 0 \\ 1 - \varepsilon_1 & \varepsilon_1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 - \varepsilon_2 & \varepsilon_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 - \varepsilon & \varepsilon \end{pmatrix}$$

mit $\varepsilon = \varepsilon_1 \varepsilon_2$. Gesamtkapazität berechnet sich nach derselben Formel

$$C_{\text{ges}} = \log_2(1 + \varepsilon(1 - \varepsilon)^{\frac{1-\varepsilon}{\varepsilon}})$$

und

$$p_X = \frac{1}{(1 - \varepsilon)^{\varepsilon-1} \varepsilon + \varepsilon}.$$

Der Gesamtkanal ist wieder ein Z-Kanal mit Parameter $\varepsilon = \varepsilon_1 \varepsilon_2$.

26.06.14

Übung 1.4.

Lösung.

(a) Der Erwartungswert der Zufallsgröße ist 0, da die Verteilung symmetrisch um 0 ist, d.h. $f(x) = \frac{1}{a} \left(1 - \left|\frac{x}{a}\right|\right)$. Die Varianz bestimmen wir, indem wir $E((X - E(X))^2)$ berechnen, dann erhalten wir also

$$\int_{-a}^a \frac{x^2}{a} \left(1 - \left|\frac{x}{a}\right|\right) dx = 2a^2 \int_0^1 y^2(1 - y) dy = \frac{a^2}{6}.$$

(b) Gegeben ist eine Zufallsgröße X mit und $Y = g(X)$ mit f_X Dichte von X und f_Y Dichte von Y . Es sei $g(x) = \exp(x/a)$. Der Transformationssatz für Dichten besagt:

$$f_Y(y) = \frac{f_X(g^{-1}(y))}{|g'(g^{-1}(y))|}$$

Ableitung ist dann $g'(x) = \frac{1}{a} \exp(x/a)$, Umkehrfunktion $g^{-1}(y) = a \log(y)$ ($y \in (0, \infty)$). Träger von Y ist $\overline{g(-a, a)} = [e^{-1}, e]$. Also für $y \in [e^{-1}, e]$ gilt:

$$f_Y(y) = \frac{\frac{1}{a} \left(1 - \left|\frac{a \log(y)}{a}\right|\right)}{\frac{1}{a} \exp\left(\frac{a \log(y)}{a}\right)}$$

insgesamt

$$f_Y(y) = \begin{cases} \frac{1}{y} (1 - |\log(y)|) & : y \in [e^{-1}, e] \\ 0 & : \text{sonst} \end{cases}$$

(c) Es ist

$$\begin{aligned} h(X) &= - \int_{-\infty}^{\infty} \log_2(f_X(x)) f_X(x) \, dx \\ &= - \int_{-a}^a \frac{1}{a} \left(1 - \left|\frac{x}{a}\right|\right) \log_2\left(\frac{1}{a} \left(1 - \left|\frac{x}{a}\right|\right)\right) \, dx \\ &= 2a^2 \int_{y=1/a}^0 y \log_2(y) \, dy \\ &= \frac{-2a^2}{\log(2)} \left(y^2 \left(\frac{\log(y)}{4} - \frac{1}{4} \right) \right) \Big|_{1/a}^0 \\ &= \log_2(\sqrt{e}a) \end{aligned}$$

Lösung. Die Dichten von X und Y sind wie folgt gegeben durch

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left(-\left(\frac{x}{\sigma_X}\right)^2\right)$$

und f_Y analog. Dann ist

$$\begin{aligned} D(X||Y) &= \int_{-\infty}^{\infty} f_X(x) \log_2\left(\frac{f_X(x)}{f_Y(x)}\right) \, dx \\ &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left(-\frac{1}{2}\left(\frac{x}{\sigma_X}\right)^2\right) \log_2\left(\sqrt{\frac{\sigma_Y^2}{\sigma_X^2}} \exp\left(-\frac{1}{2}\left(\frac{1}{\sigma_X^2} - \frac{1}{\sigma_Y^2}\right)x^2\right)\right) \, dx \\ &= \frac{1}{2} \log_2\left(\frac{\sigma_Y^2}{\sigma_X^2}\right) E(X) + \frac{1}{\log(2)} \frac{1}{2} \left(\frac{1}{\sigma_Y^2} - \frac{1}{\sigma_X^2}\right) V(X) \\ &= \frac{1}{2 \log(2)} \left(2 \log\left(\frac{\sigma_Y}{\sigma_X}\right) + \left(\frac{\sigma_X^2}{\sigma_Y^2} - 1\right)\right) \end{aligned}$$

Lösung. Wir haben $Z_1 := X_1 + Y$, $Z_2 := X_2 + Y$, Z_1 ist normalverteilt, denn die Summe normalverteilter Zufallsgrößen ist normalverteilt. $E(Z_1) = E(X_1) + E(Y) = 0 + 0 = 0$ und die Varianz ist $V(Z_1) = V(X_1) + V(Y) = \sigma_1^2 + \sigma_Y^2$, da X_1, Y unkorreliert (da unabhängig). Entsprechend ist $E(Z_2) = 0$ und $V(Z_2) = \sigma_2^2 + \sigma_Y^2$.

(a) Mit Aufgabe 32 erhalten wir:

$$D(X_1 + Y || X_2 + Y) = \frac{1}{2 \log(2)} \left(\log\left(\frac{\sigma_2^2 + \sigma_Y^2}{\sigma_1^2 + \sigma_Y^2}\right) + \frac{\sigma_1^2 + \sigma_Y^2}{\sigma_2^2 + \sigma_Y^2} - 1 \right)$$

(b) Es gilt $D(X_1||X_2) \geq D(X_1 + Y||X_2 + Y)$, da

$$\log\left(\frac{b+x}{a+x}\right) + \frac{a+x}{b+x} - 1$$

monoton fallend in x ist. Die Ableitung ist

$$\frac{a+x}{b+x} \frac{a-b}{(a+x)^2} + \frac{b-a}{(b+x)^2} = \frac{a-b}{b+x} \left(\frac{1}{a+x} - \frac{1}{b+x} \right) = \frac{-(b-a)^2}{(b+x)^2(a+x)} < 0$$

Bemerkung 4. Durch additive Störung werden die Verteilungen der Zufallsgrößen X_1 und X_2 ähnlicher, daher $D(\bullet||\bullet)$ kleiner.

Lösung. Äquivalentes Modell: $\tilde{Z} = Z_1 + Z_2$, \tilde{Z} normalverteilt, und $Y = X + \tilde{Z}$.

(a) $E(\tilde{Z}) = E(Z_1) + E(Z_2) = 0$ und $V(\tilde{Z}) = V(Z_1) + V(Z_2) = \sigma_1^2 + \sigma_2^2$ (da Z_1 und Z_2 unabhängig sind. X und (Z_1, Z_2) sind unabhängig, also Modell GAUSS-Kanal (Hintereinanderschaltung von zwei GAUSS-Kanälen).

$$\max_{E(X^2) \leq P} I(X; Y) = \max_{E(X^2) \leq P} I(X; X + Z) = \frac{1}{2} \log_2 \left(1 + \frac{P}{\sigma_1^2 + \sigma_2^2} \right)$$

nach der Formel für die Informationskapazität des GAUSS-Kanals (§3.51).

(b) Es ergibt sich

$$\begin{aligned} V(\tilde{Z}) &= V(Z_1) + 2C(Z_1, Z_2) + V(Z_2) \\ &= \sigma_1^2 + 2\alpha\sigma_1\sigma_2 + \sigma_2^2 \end{aligned}$$

Wieder muss in die Formel für die Transinformation eingesetzt werden.